**Diagnosing Bias vs. Variance**

In this section we examine the relationship between the degree of the polynomial d and the underfitting or overfitting of our hypothesis.

* We need to distinguish whether **bias** or **variance** is the problem contributing to bad predictions.
* High bias is underfitting and high variance is overfitting. Ideally, we need to find a golden mean between these two.

The training error will tend to **decrease** as we increase the degree d of the polynomial.

At the same time, the cross validation error will tend to **decrease** as we increase d up to a point, and then it will **increase** as d is increased, forming a convex curve.

**High bias (underfitting)**: both Jtrain(Θ) and JCV(Θ) will be high. Also, JCV(Θ) ≈ Jtrain(Θ).

**High variance (overfitting)**: Jtrain(Θ) will be low and JCV(Θ) will be much greater than Jtrain(Θ).

The is summarized in the figure below:

